

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | | |
|-------------------------------------|--|
| n/a | Confirmed |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of all covariates tested |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Cells were imaged with either a Nikon Ti-E or Nikon Ti2 fluorescence microscope and controlled by Micro-Manager or Nikon Elements. Data was processed using the scientific computing stack for python. The requirements, including version and usage information, are available at <http://www.github.com/vanvalenlab>

Data analysis

We used Kubernetes and TensorFlow, along with the scientific computing stack for python. A persistent deployment of the software described can be accessed at <http://www.deepcell.org>. All source code, including version requirements and explicit usage, is under a modified Apache license and is available at <http://www.github.com/vanvalenlab>. Detailed instructions are available at <http://deepcell.readthedocs.io/>.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All data that were used to generate the figures in this paper are available at <http://www.deepcell.org/data> and at <http://github.com/vanvalenlab/deepcell-tf> under the deepcell.datasets module.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Given the inherent lack of variability in our computational system, no sample-size calculation was performed. Rather, in our benchmarking, we sought to select sample sizes and data rates comparable to the average imaging experiment (100 images corresponds to a 96-well plate, common ethernet connections setting data rates, etc.). Benchmarking on runs of 10,000 to 1,000,000 images was selected to satisfy the upper limit of analysis needs. We believe these sample sizes are sufficient to illustrate the effectiveness of this massively parallel computing scheme. Each benchmarking experiment (from 10,000 to 1,000,000) was run in triplicate with the exception of the 1,000,000 image case due to cost considerations.
Data exclusions	No data was excluded from the analyses.
Replication	The data, algorithms, and software described have a manuscript release version to ensure reproducibility All data that were used to generate the figures in this paper are available at http://www.deepcell.org/data and at http://github.com/vanvalenlab/deepcell-tf under the deepcell.datasets module. The version of software that produced the results reported is permanently available at: https://github.com/vanvalenlab/kiosk/releases/tag/1.0.0 .
Randomization	Randomization is not relevant to this study. The system has some inherent randomness, insomuch as the physical machines that are selected to process any given image may be different each time, but the only additional concern would be the content of the images in each benchmarking run. These were selected and duplicated from existing data to match the image needs of the run in question and the segmentation model used. Similar source data was used for each run to ensure a valid comparison.
Blinding	Blinding was not relevant to this study because there is no risk of bias in this computational system.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

Methods

n/a	Involved in the study	n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies	<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines	<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology	<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data		

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	We used the mammalian cell lines NIH-3T3, HeLa-S3, HEK 293, and RAW 264.7 to collect training data for nuclear segmentation and the cell lines NIH-3T3 and RAW 264.7 to collect training data for augmented microscopy. All cell lines were acquired from ATCC.
Authentication	The cells have not been authenticated.
Mycoplasma contamination	The cells were not tested for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	N/A